

Data management pilot project report

Report date	2017 06 15		Report #	2
Project title	ScienceCloud4Cities			
Grant holder	Assoc. Prof. Alfred Heller, DTU			
Partner institutions	SDU(b) + AU			
Project start	1/8-2016	Project end	1/5-2018 (org. 02/12.2017)	

Details about the reporting

Video conference June 15th 2017, 14.30 – 15.30
Participants: Project side: PI Alfred Heller. Evaluator side: Henrik Pedersen + Rene Belsø + Josva Kleist + Diba Markus

Overall assessment at this point in time

The project has a success already as it used extensively by Cities researchers.

Action points for grant holder, DM Secretariat or others

	Assignee	Deadlines	Who's in the loop
Henrik has to bring the question of the lacking co-financing before DeiC CEO Steen Pedersen (ultimately it may be a board decision)	Henrik	ASAP	Alfred
PI (still) needs help with tackling the increased demands of the GDPR	Henrik	ASAP	Alfred

Project progress so far

Projektet har arbejdet på en aftale mellem partnerne i projektet og der mangler kun DTUs egen underskrift, hvilket afhænger af en intern aftale der er besværet af den høje medfinansieringsbidrag for projektet. Situationen er beskrevet af mit overfor Deic gennem de seneste måneder og gentages ikke her.

Et andet kritisk emne er nævnt sidste gang og handler om EU's lovramme for data. Her er udviklingen ikke kommet videre, hverken i CITIES eller Nordhavn projektet, mens der i Smart Cities Accelerator projektet - som er et EU projekt med deltagere fra Sverige og Danmark og hvor SC4C er cloudkomponenten for alle forskningsaktiviteter fra indeklima i skoler, over fjernvarme til brugerinddragelse i offentlige projekter – er inddraget juridiske eksperter fra hhv. KU og Lund Universitet for at finde ud af disse forhold.

Projektet er meget succesfuld i forhold til de forsknings- og studiemæssige resultater der er opnået:

- flere studenterprojekter har bygget på infrastrukturer allerede, disse er fra hhv. DTU og AU.
 - Vi kan speciel fremhæve et projekt hvor sensitive data bliver anonymiseret ved "analyse" (ikke aggregering eller andet anonymisering), hvor der findes dataenes statistiske egenskaber som publiceres, samt suppleres ved syntetiske data fra simuleringer i avancerede modeller. Projektet publicerer disse data i egen repository som skal hjælpe studerende og forskere at finde typiske bygningsenergidata for danske bygninger. Dataene bliver herudover publiceret på de internationale platforme Zenodo og den nationale opendata.dk (CKan) – metoderne er stadig under afprøvning – projektet bliver forsvaret i juli i år.
 - Andre studenterprojekter bruger data til at bygge styringsteknologier som MPC'er.
- Mange PhD og postdoc fra bl.a. CITIES og sammenhængende projekter, bygger på infrastrukturen og indeholder såvel sensitive som ikke sensitive eksempler på data håndtering.
 - Vi kan fremhæve et eksempel med data fra forbrug af elektricitet i område Sønderjylland som er ekstremt stort og har givet udfordringer med denne størrelse. Vi er ved at lægge sidste hånd på overførelse af dataene til vores infrastruktur og forventer denne sag afsluttet inden sommer, inkl. håndtering af aftaleramme for sådanne data, hvor vi henter inspiration fra Zenodo.
- Flere nationale og internationale forskningsprojekter anvender infrastrukturen: Her kan nævnes CITIES, Smart Cities Accelerator, Carbon Track and Traces, samarbejdet i Vidensby Lyngby og mange flere. Også er infrastrukturen foreslået i det store nationale forskningsprojekt REBUS som har deltagere fra flere universiteter, byer og virksomheder.

Challenges

- En endelig "arkitektur" for integration med de mange eksterne data kilder og data aftager infrastrukturer – der er mange muligheder og vi skal finde vores vej at anvende disse på – f.eks. hvornår skal man lægge data på Zenodo og hvornår ønsker vi dem på vores infrastruktur og kun dele metadata internationalt?
- Der mangler aktivering af nogle partnere som vi forventer ske i efterår 2017 og forår 2018.
- Vi har nedprioriteret integration med HPC fra SDU pga. gensidig lav kommunikation – vi har ikke tænkt os at prioritere denne indsats op.
- Omvendt vores prioritering af samarbejde med de andre universitetspartnere bliver opprioriteret i den kommende periode da vi arbejder på CITIES afløser som kommer til at hedde Cities Innovation Centre – her kommer SC4C til at spille en afgørende rolle som den allerede gør i andre projekter der er nævnt her.
- Vi arbejder på en "data publicerings case" for at vise forskere og studerende, hvordan man gør dette med vores platform – også her afprøver vi muligheder og har ikke fundet en endelig metode endnu. Studentercasen nævnt ovenfor er en helt oplagt case til dette da den er en proof of concept til vores IEA samarbejder der er nævnt andet steds.

Ekstra indsatser som giver omtale af DeiC og SC4C

Der kan diskuteres om man i projektet kan kompenseres for enkelte leverancer ved at der vises nogle andre resultater:

- Ved High Tech Summit på DTU i September 2017 kommer SC4C til at være en af de resultater som DTU kan vise frem til sine partnere – her bliver Smart Cities tracken, IoT tracken og Smart Bygning tracken muligheder at promovere løsningen (DeiC kunne overveje at blive deltager og bidrage med egne resultater – data.deic.dk).
- SC4C er anvendt og foreslået til mange nationale og internationale forskningsprojekter der overgår i størrelse alt det man kunne drømme om.
- DeiC bevillingen er fremført i konference publikationer og journal publiceringer.
- Da andre projekter ikke kan levere en data infrastruktur (med mindre man går efter en kommerciel løsning), så bliver vores infrastruktur omtalt og prioriteret i rigtig mange sammenhænge – enkelte vælger vores system til deres udvikling. Forskere er meget glad for vores platform da den opleves som en "lokal maskine med kendt operativsystem" overfor dem da man kun ved login mærker at man er på en cloud service.

Up for discussion:

- Tiple helix adgang skulle muliggøre for at give alle adgang til smart city data – Den Europæiske data lovgivning gør at vi vælger ikke at dele data fra forskerene tilbage til f.eks. byer og industripartnere uden ekstra aftaler. Dermed er denne integration ikke relevant før en afgørelse er sket.
- Integration med data.deic.dk er allerede situationen da data.deic.dk bliver brugt som "sharing" portal for data, dog kunne en integration sker meget mere omfattende.
- Der mangler en endelig afklaring på metadata område som SDU står for – denne aktivitet kræver en endelig aftale på DTU internt – se ovenfor.
- Real time data casen er fravalgt af to grunde, da denne i CITIES sammenhænge bliver leveret af en anden komponent SE-OS fra DTU Compute og da der ikke var en case der skulle håndteres denne vej (Selv data tracking casen i Lyngby kræver ikke så aktuelle data).